Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 6, No. 1, 2023

Power Quality Disturbance Recognition based on Deep Neural Network and Adaptive Feature Fusion

Lei Chen, Chao Zhou *, Jianjun Chen

Department of Telecommunications, Qinhuangdao Campus, Northeast Petroleum University, Qinhuangdao, Hebei 066044, China
*** Corresponding author:** Chao Zhou (Email: 1239483277@qq.com)

Abstract: Aiming at the problems of low recognition accuracy and low noise resistance of current power quality disturbance recognition algorithms, a power quality disturbance recognition method based on the fusion of neural network and adaptive features is proposed. Firstly, one-dimensional features are extracted by 1D_CNN+GRU network. The GADF algorithm and 2D_CNN are used to extract 2D features, and then the extracted 1D features and 2D features are adaptively fused into a new feature. Finally, the new features are input into the channel attention mechanism and classified through the full connection layer. The experimental results show that the classification accuracy of the proposed method is above 92% for 40 disturbance types containing single, double, triple and quadruple disturbances, and above 96% for 20 disturbance types containing single and double disturbances.

Keywords: Power Quality Disturbance; Neural Network; Feature Fusion; Adaptive Weights; GADF Algorithm.

1. Introduction

With the increasing of nonlinear load and the rapid development of smart grid, power enterprises and users pay more and more attention to power quality[1]. In the power system, after the power quality is disturbed, the voltage signal and the current signal will change [2] in different forms, so that the electronic equipment is in the abnormal working state. The identification and classification of power quality disturbance is an effective method [3]to control and improve power quality.

The problem of power quality disturbance identification and classification is generally divided into two parts, that is, feature extraction and classifier design. Usually, feature extraction is carried out first, and then a model is established to classify the disturbance. The existing feature extraction methods are: Fourier transform[4], wavelet transform, S transform[5][6], Hilbert-Yellow transform[7] and so on. Disturbance classification is based on feature extraction, combined with different classification methods, to determine the category of disturbance. The commonly used classifiers are: decision tree[8], artificial neural network[9] and support vector machine[10]. Literature [11] uses S-transform to extract the characteristics of power quality disturbance data and chaotic integrated decision tree algorithm to classify the disturbance. Compared with traditional decision tree method, the recognition accuracy of this method is greatly improved. In literature [12], mRMR algorithm is used to extract the features of power quality disturbance data, and then multi-core SVM is used to classify the extracted disturbance features. This method improves the operation speed, has strong anti-noise ability and good stability. Literature [13] uses improved S-transform (SMST) to extract the time-frequency features of the power quality disturbance signals to be measured, and uses the improved CART algorithm to build a random forest (RF) classifier to classify the disturbance features. Compared with the traditional CART algorithm, this method has a great improvement in the identification accuracy of single and double compound disturbances. In literature [14], empirical wavelet transform (EWT) is used to obtain the eigenmatrix of power quality disturbance signals, and the obtained eigenmatrix is input to support vector machine (SVM) for disturbance classification. EWT is built on the basis of wavelet transform (WT). Compared with WT, EWT has the advantage of empirical mode decomposition (EMD), and the calculation speed is faster. The author simulated the power quality disturbance model under different conditions, and the method showed good performance for single disturbance identification accuracy and noise resistance.

The above feature extraction methods all require manual intervention, and the features determined manually are completely dependent on human experience, which is difficult to fully reflect the key features of the disturbance, resulting in reduced accuracy of disturbance identification and weak adaptability. The deep learning method can eliminate the constraint of human factors and carry out intelligent feature extraction from a large number of disturbance data. It can adapt to the needs of PQ disturbance intelligent classification in the current era of big data. In this paper, a power quality disturbance recognition method based on adaptive feature fusion is proposed, which realizes the recognition and classification of power quality disturbance data by adaptive fusion of one-dimensional and two-dimensional features extracted from two subnetworks.

2. PQDs Signal Generation and Processing

2.1. Gramian Angular Field (GAF)

The Gram Angle field algorithm can realize the reconstruction of one-dimensional time series, and can avoid the data loss in the reconstruction process. Compared with other conversion methods, GAF polar coordinates retain the absolute value of time relation, and the original time series in Cartesian coordinates can be restored through the main diagonal of GAF matrix.

The principle of Gram angular field algorithm is as follows: firstly, the value scaling is carried out; $X = \{x_1, x_2, \cdots, x_n\}$The time series in the Cartesian coordinate

system is scaled to the interval [1, -1] by formula (1), and then the polar coordinates are converted; Formula (2) will be converted to polar coordinate system time series.$X\widetilde{X}$

$$\tilde{x}_i = \frac{x_i - \max X + x_i - \min X}{\max X - \min X} \tag{1}$$

$$\begin{cases} \theta = \arccos \tilde{x}_i \ ; \ -1 \le \tilde{x}_i \le 1, \tilde{x}_i \in \tilde{X} \\ r = \dfrac{t_i}{N} \ ; \ t_i \in N \end{cases} \tag{2}$$

Where: is the original time series; $x_i\tilde{x}_i$Is the time series after reconstruction; t_iIs the time stamp; NAnd the constant factor of the generated space for the regularized polar coordinate system.

If the cosine function of the sum of two angles is used, the Gram Angle and field (GASF) diagram can be obtained, and if the sine function of the difference of two angles is used, the Gram Angle difference field (GADF) diagram can be obtained. Compared with GASF, GADF has a better [17] ability to express the features of the original disturbed data. In this paper, the feature map of GADF is adopted. GADF calculation method is as follows:

$$G_{GADF} = \begin{pmatrix} \sin(\theta_1 - \theta_1) & \cdots & \sin(\theta_1 - \theta_n) \\ \vdots & \ddots & \vdots \\ \sin(\theta_n - \theta_1) & \cdots & \sin(\theta_n - \theta_n) \end{pmatrix} = \sqrt{I - \left(\tilde{X}^2\right)^T} \bullet \tilde{X} - \tilde{X}^T \bullet \sqrt{I - \tilde{X}^2} \tag{3}$$

Where: is the unit row vector; $I\tilde{X}^T$Is the transpose vector of.$\tilde{X}$

2.2. Data Generation and Processing

In this paper, the original one-dimensional disturbance data are generated by numerical calculation software MATLAB, including a total of 40 types of power quality disturbance data, including 8 types of single disturbance, 15 types of double disturbance, 13 types of triple disturbance and 4 types of quadruple disturbance, as shown in Table 1. The original one-dimensional data is converted into two-dimensional data by GADF algorithm. FIG. 1 shows the GADF diagram of 8 common single disturbance and normal waveforms, in which the disturbance types from left to right in the first row are interrupt (C3), pulse (C6), dip + harmonic (C8), and rise + oscillation (C15). In the second row, the disturbance types from left to right are respectively sag + oscillation + pulse (C24), harmonics + interrupt + pulse (C29), harmonics + oscillation + pulse + temporary rise (C37), harmonics + oscillation + wave + interrupt (C39).

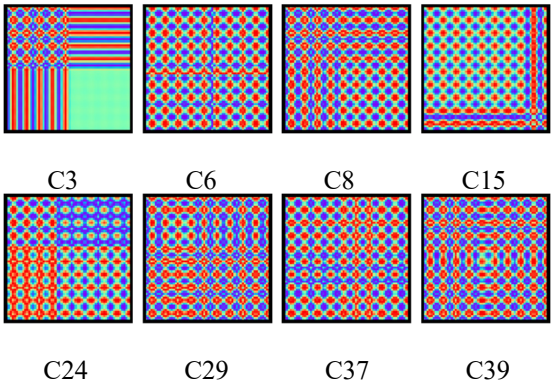


Fig 1. Partial disturbance GADF diagram

Table 1. Perturbation types correspond to the symbols in this document

Text symbols	Type of disturbance
C0	Normal waveform
C1	Sagging
C2	Temporary promotion
C3	Interrupt
C4	Harmonics
C5	Oscillations
C6	Pulse
C7	Undulation
C8	Dip + harmonics
C9	Temporary rise + harmonics
C10	Interrupt + harmonics
C11	Oscillations + harmonics
C12	Pulse + Harmonics
C13	Wave + Harmonics
C14	Dip + Oscillation
C15	Temporary rise + Oscillation
C16	Interrupt + Oscillation
C17	Pulse + Oscillation
C18	Wave + oscillation
C19	Volatility + temporary decline
C20	Volatility + temporary rise
C21	Volatility + Disruption
C22	Wave + Pulse
C23	Harmonics + oscillations + pulses
C24	Dip + oscillation + pulse
C25	Temporary rise + Oscillation +
C26	Interrupt + oscillation + pulse
C27	Harmonics + dip + pulse
C28	Harmonics + temporary rise +
C29	Harmonics + interrupt + pulse
C30	Harmonics + oscillations + dips
C31	Harmonics + oscillations +
C32	Harmonics + oscillation + interrupt
C33	Harmonics + Fluctuations +
C34	Harmonic + wave + temporary rise
C35	Harmonics + Wave + dip
C36	Harmonic + Oscillation + pulse +
C37	Harmonic + Oscillation + pulse +
C38	Harmonics + oscillations + pulses
C39	Harmonics + oscillations + waves

3. Neural Network

3.1. Convolutional Neural Networks (CNN)

The types of CNNS used in this paper are 1D_CNN and 2D_CNN. The CNNS can be divided into convolutional layer, pooling layer, activation function, batch normalization layer and fully connected layer.

3.1.1. Convolutional Layer

The convolution kernel in the convolution layer is the computing unit. The convolution kernel moves on the feature map of the input convolution layer and carries out convolution operations on each region of the feature map. The calculation results of different convolution check feature maps are different. The result of the convolution plus the offset of the output feature map as the output of the convolutional layer. The convolutional layer is calculated as follows:

$$x_j^l = \sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_i^j \quad (4)$$

Where: represents the number of layers; lM_jRepresents the set of feature maps connected with the current layer's first feature map in the previous layer; jx_j^lIs the first feature map output of the first layer; ljx_i^{l-1}The first feature map output for the first layer; $l-1ik_{ij}^l$Is the convolution kernel between the first feature map of the first layer and the first feature map of the previous layer; $ljib_j^l$Is the bias of the first layer's first feature map.lj

3.1.2. Pool Layer

The pooling layer is located after the convolution layer. Since the convolution layer increases the number and dimension of feature maps, in order to reduce the number and dimension of feature maps, and keep the scale invariance of features as much as possible. The pooling layer can extract the features twice after the convolution layer, which improves the generalization ability of the model and is not easy to overfit. The calculation process of the pooling layer is as follows:

$$x_j^l = \beta_j^l subdown\left(x_i^{l-1}\right) + b_j^l \quad (5)$$

Where: represents the pooled downsampling function; $subdown(\cdot)\beta_j^l$Is proportional bias; b_j^lIs additive bias.

3.1.3. Batch Normalization Layer (BN)

BN layer can overcome the influence of covariance deviation, speed up the training speed, and enhance the generalization ability of the network.

3.1.4. Activate Function

In order to avoid the problem of insufficient expressiveness of the neural network model, a nonlinear function is used as the activation function of the feature graph output by the pooling layer. Commonly used activation functions include Sigmoid activation function, TanH activation function, ReLU activation function, etc. Since the first two activation functions belong to saturation activation function, gradient disappearance will occur due to their saturation, and the convergence speed is slow, so this paper chooses the unsaturated activation function ReLU activation function, which has a faster convergence speed. ReLU activation function is defined as follows:

$$f(x) = \begin{cases} 0 & (x<0) \\ x & (x \geq 0) \end{cases} \quad (6)$$

3.1.5. Full Connection Layer

After multiple convolution, pooling and activation operations of the input data, the resulting feature is expanded into a one-dimensional vector, so that the feature corresponds to the disturbance type one by one, expressed as:

$$y^k = f\left(w^k x^{k-1} + b^k\right) \quad (7)$$

Where: is the number of layers of the network; ky^kIs the total output of this layer; x^{k-1}Is the initial input of this layer; w^kIs the weight; b^kFor network bias.

3.2. Gated Cycle Unit (GRU)

Grus were proposed to solve problems such as long-term memory and gradients in backpropagation. The GRU is composed of two gates: update gate and reset gate. The update gate controls the influence of the output of the hidden layer at the previous time on the output of the hidden layer at the same time, and the reset gate indicates the information output of the hidden layer at the previous time to be deleted at the same time, which can effectively alleviate the problem [16] of "gradient disappearance" in the recurrent neural network.

4. Model Establishment and Simulation Experiment

4.1. Establishment of Adaptive Feature Fusion Model

4.1.1. 1D_CNN+GRU

The module consists of 1D_CNN combined with GRU. 1D_CNN consists of 5 convolutional layers, 5 normalized layers, and no pooling layer. The size of the convolutional kernel is 2×2, the move step of the convolutional kernel is set to 2 and no filling is required, and the parameters of dropout are set to 0.5. The one-dimensional data is taken as the input of 1D_CNN, and then the CNN output is taken as the input of GRU, and the output of GRU is the one-dimensional data feature. The 1D_CNN+GRU structure is shown in Figure 2.

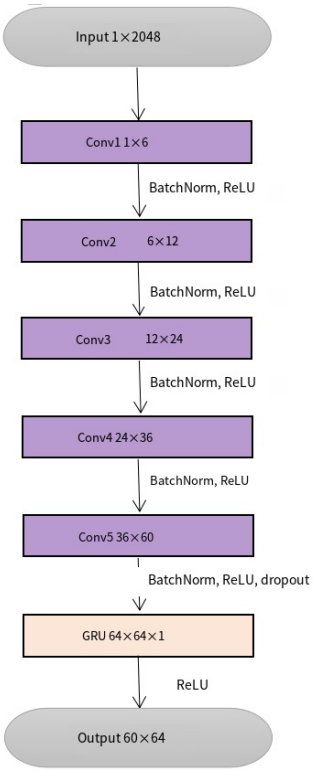


Fig 2. 1D_CNN+GRU Structure Diagram

4.1.2. 2D_CNN

This module consists of 2D_CNN, obtained by modifying the classical neural network model structure VGG16. This module consists of 6 convolutional layers, 3 pooled layers and 6 normalization layers. The convolutional kernel size of the convolutional layer is 3×3, the convolutional kernel moving step is 1, the picture filling is 1, the pooled window of the pooled layer is 2×2, and the pooling window moving step is 2. Through the above parameter Settings, the size of the input through the convolutional layer remains unchanged, and the size is changed only through the pooled layer. The parameters of dropout are set to 0.5, and finally the output is activated by the ReLU function. The 2D data is input as this module, the output shape is (64,60,8,8), and the output shape is changed to (64,60,64) by changing the array shape function, as the 2D data feature. The 2D_CNN structure is shown in Figure 3.

Fig 3. 2D_CNN Structure Diagram

4.1.3. Adaptive Feature Fusion

In order to make better use of the extracted one-dimensional and two-dimensional feature vectors, adaptive feature weights are introduced to make the model allocate different weights according to the feature distribution to carry out feature fusion [17]. The structure of this module is shown in Figure 4. The one-dimensional and two-dimensional features are multiplied by w1 and w2 respectively and then added, and the added result is taken as the input of SENet, and finally the fused feature vector is output. The initial weights w1 and w2 are initialized by Parameter function, both of which are 0.5. In the network, the initial weight parameter will be added to the update parameter of the optimizer, so that when the optimizer is updated, the optimization direction is to minimize the loss function. SENet is the channel attention module, which can obtain the importance degree of each channel in the feature graph and automatically assign different weights according to the importance degree. The scope of adaptive feature weights is the overall operation of one-dimensional and two-dimensional features, while the scope of SENet is the internal channel of the feature vector.

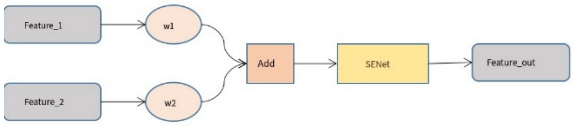


Fig 4. Structure diagram of adaptive feature fusion module

4.1.4. Overall Flow Diagram

In this paper, the error backpropagation method is used to train the convolutional neural network. The training samples pass through the convolutional layer, the pooling layer, the activation function activation and the full connection layer, and the weight and bias of each layer are updated by the random gradient descent algorithm. The overall process is shown in FIG. 5. One-dimensional data is extracted by 1D_CNN+GRU module with one-dimensional features; One-dimensional data is converted into two-dimensional data by GADF algorithm, two-dimensional features are extracted by 2D_CNN module, and feature fusion is carried out by adaptive feature fusion module. After fusion, classification is output through the full connection layer.

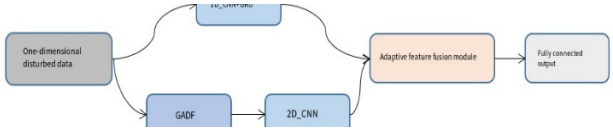


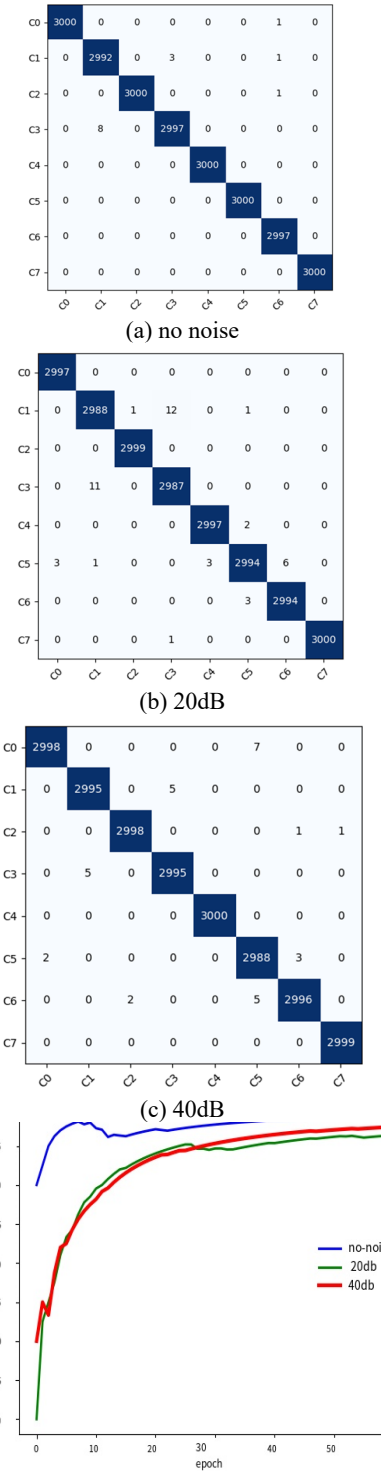
Fig 5. Overall flowchart

4.2. Training and Testing of the Model

In this paper, 500 pieces of data were selected for each disturbance, with a total of 20,000 pieces of data for 40 kinds of disturbances. 2048 data points were collected for each disturbance by Matlab with a total sampling time of 0.2s. The original data was divided into training set, verification set and test set according to the ratio of 8:1:1. In order to test the anti-noise ability of the model, white noise is superimposed on the disturbance signal, and the signal-to-noise ratio is 20dB and 40dB respectively. The model was trained and tested by three mixed modes: single disturbance (T1), single disturbance and double disturbance (T2), and multiple compound disturbance (T3) with single, double, triple and quadruple disturbance.

60 cycles were set for the training and testing of this model. During the training process, Gaussian noise or salt and pepper noise were added to the data respectively to achieve data enhancement. The loss function was cross-entropy function and the optimizer was Adam. The learning rate was reduced from 0.01 to 0.003 by exponential decay learning rate function ExponentialLR function. Figure 7 to 9 shows the training process of T1, T2 and T3 with no noise and signal-to-noise ratio of 20dB and 40dB respectively.

4.2.1. Classification of T1



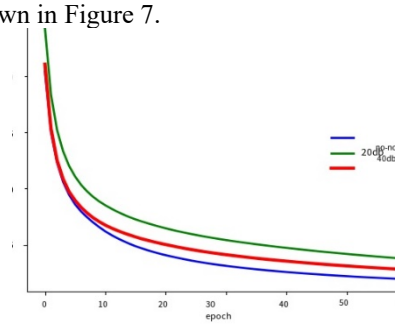
(d) validation-acc curve with different signal-to-noise ratio

Fig 6. Training process of 8 classification with different signal-to-noise ratios

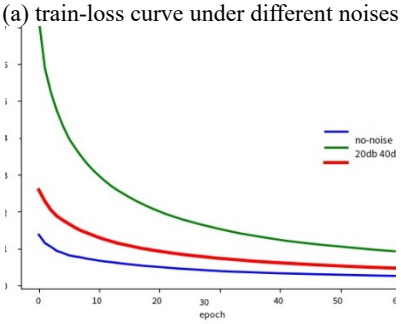
The confusion matrix can intuitively observe the classification of the model. The horizontal coordinate of the matrix is the type predicted by the model, and the vertical coordinate is the real category, in which the diagonal elements are the number of correct classification, and the non-diagonal elements are wrong judgment and missing judgment. Figure 6 shows the confusion matrix and training curve of the model for T1 when there is no noise and the SNR is 20dB and 40dB respectively. It can be seen from the matrix that at 20dB signal-to-noise ratio, the number of false judgments and missed judgments in the proposed method is low, especially the recognition accuracy of temporary rise and fluctuation is the highest, which are 99.96% and 100% respectively.

4.2.2. Classification of T2

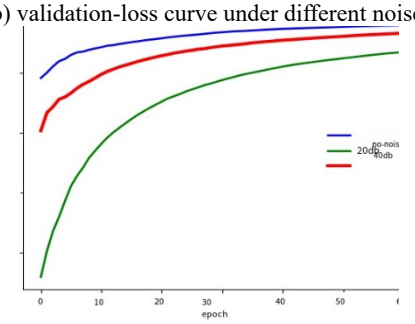
The first 20 types of disturbance in this paper include single disturbance and double disturbance. The training process of classifying the first 20 types of disturbance under different SNR is shown in Figure 7.



(a) train-loss curve under different noises



(b) validation-loss curve under different noises



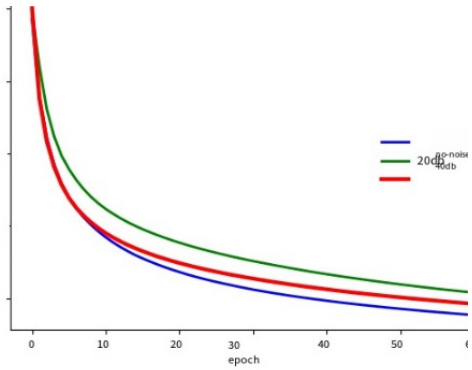
(c) validation-acc curve under different noises

Fig 7. Training process of 20 classification with different signal-to-noise ratios

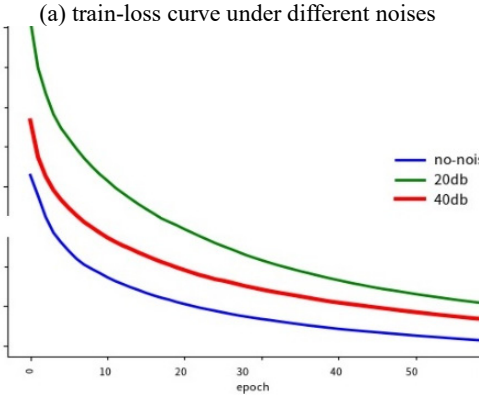
4.2.3. Classification of T3

This paper has a total of 40 types of disturbance data, including a variety of compound disturbances including single, double, triple and quadruple disturbances. The training process of classifying 40 types of disturbances under different signal-to-noise ratio is shown in Figure 8.

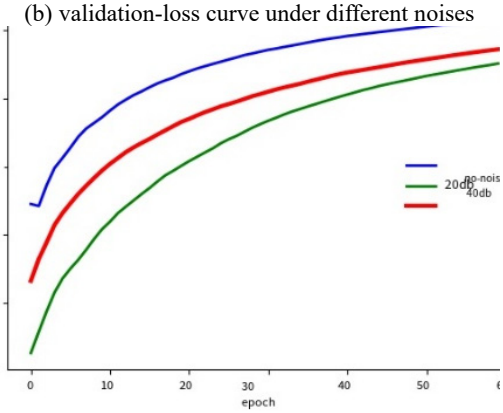
Table 2 shows the average recognition accuracy of the model in this paper for 8 classification with different signal, 20 classification and 40 classification under three different noise environments.



(a) train-loss curve under different noises



(b) validation-loss curve under different noises



(c) validation-acc curve under different noises

Fig 8. Training process of 40 classification with different signal-to-noise ratios

Table 2. The classification accuracy of the method in this paper

Number of disturbance types	Average recognition accuracy %		
	Noise-	20dB	40dB
8	99. 94	99. 81	99. 87
20	99. 00	96. 73	98. 32
40	95. 80	92. 57	93. 62

As can be seen from the above table, this model can effectively identify power quality disturbance signals regardless of 8 classification, 20 classification or 40 classification. Under the three different noise environments, the average accuracy of 8 classification is higher than 99%, the average accuracy of 20 classification is higher than 96%, and the accuracy of 40 classification is higher than 92%.

Table 3 shows the classification results of 20 types of disturbances under different signal-to-noise ratios.

As can be seen from the above table, the proposed model can effectively identify T1 and T2, and the average recognition accuracy is above 96% under three different SNR conditions.

Table 3. The classification accuracy of this method 20 under different signal-to-noise ratios

Type of disturbance	Recognition accuracy %		
	No noise	20dB	40dB
C0	100. 00	99. 26	99. 96
C1	99.66	91. 33	96. 76
C2	99.56	95. 43	96. 83
C3	99. 53	97. 33	97. 73
C4	99. 86	95. 03	99. 76
C5	99. 86	99. 26	99. 70
C6	99. 83	99. 20	99. 70
C7	100. 00	99. 96	100. 00
C8	92. 80	93. 43	95. 63
C9	99. 86	98. 56	99. 60
C10	93.46	95. 16	95. 10
C11	99. 83	94. 73	98. 60
C12	99. 90	98. 16	99. 50
C13	100. 00	98. 96	99. 93
C14	98.40	94. 16	95. 46
C15	99. 40	95. 46	97. 13
C16	98. 53	93. 30	95. 80
C17	99. 86	99. 56	99. 83
C18	99. 86	98. 30	99. 76
C19	99. 86	97. 93	99. 70
Average	99. 00	96. 73	98. 32

4.3. Comparison of 3 Methods

The proposed method is compared with 1D_CNN+GRU, 2D_CNN and 2D_ResNet [18], respectively, and the classification accuracy of 40 kinds of multiple composite disturbances under three SNR conditions of 0dB, 20dB and 40dB are respectively used by these four methods. The results are shown in Table 4.

Table 4. Classification accuracy of the four methods

MODELS	Average recognition		
	Noise-	20dB	40dB
Method of this	95. 80	92.	93. 62
1D_CNN+GRU	91.71	83.64	91.29
2D_CNN	93.95	92.28	93.57
2D_ResNet	86.39	81.51	85.28

As can be seen from Table 4, the average recognition accuracy of the proposed method is higher than that of the other three under three different noise conditions. When there is no noise, the recognition accuracy of the proposed method is increased by 4.45%, 1.97% and 10.89% compared with the methods 1D_CNN+GRU, 2D_CNN and 2D_ResNet, respectively. When the SNR is 20dB, the recognition accuracy of the proposed method is improved by 10.67%, 0.31% and 13.56%, respectively, compared with the other three methods. When the SNR is 40dB, the recognition accuracy is improved by 2.55%, 0.05% and 9.77%, respectively. It can be seen from the first three methods that the recognition accuracy and noise resistance of the model after feature fusion are significantly improved compared with the model without feature fusion.

5. Conclusion

In order to avoid the decrease of recognition accuracy caused by manual extraction of disturbed features, this paper uses deep neural network to automatically extract features. In order to make the model have higher recognition accuracy and noise resistance, this paper proposes a deep neural network algorithm of adaptive feature fusion. The experimental results show that under three different SNR conditions, the recognition accuracy of the model for 8 types of disturbance with only one disturbance is higher than 99%. The recognition accuracy of 20 kinds of disturbances including single disturbance and double compound disturbance is greater than 96%; For 40 types of multiple complex disturbances containing single, double, triple and quadruple disturbances, the recognition accuracy is above 92%. Compared with the other three methods, the recognition accuracy and noise resistance of the proposed method are improved in different degrees, which proves the feasibility of the model and provides a new idea for dealing with today's complex types of power quality disturbances.

References

[1] Wang Weibo, Zhang Bin, Zeng Wenyu, et al. Power quality disturbance classification based on feature fusion one-dimensional Convolutional neural network [J]. Power System Protection and Control, 2020, 48(06): 53-60.

[2] Sebastijan S, Niko L, Boheme, et al. Power quality experimental analysis of grid-connected photovoltaic systems in urban distribution networks[J]. Energy, 2017, 139.

[3] Jin Guo, Zhu Qing-Zhi, Meng Yang, et al. Power Quality disturbance multi-label classification algorithm based on multi-layer extreme Learning Machine [J]. Power System Protection and Control, 2020, 48(08): 96-105.

[4] HUANG Jianming, LI Xiaoming. Harmonic detection method of power system combined with short-time Fourier transform and spectral steepness[J]. Power System Protection and Control, 2017, 45(07): 43-50.

[5] Tao Caixia, Du Xue, Gao Fengyang, et al. Fault location for single-phase grounding of hybrid transmission lines based on empirical wavelet transform [J]. Power System Protection and Control, 2021, 49(10): 105-112.

[6] Cheng Zhiyou, Yang Meng. Power quality disturbance type recognition based on two-dimensional discrete cosine S transform [J]. Power System Protection and Control, 2021, 49(17): 85-92. (in Chinese)

[7] Luo Xuelian, Liu Guiying, Liu Xiangyin. Research on Detection and Identification of transient power quality Disturbance in Microgrid based on Wavelet transform and Hilbert-Yellow Transform [J]. Power Capacitors and reactive power compensation, 2020, 41(03): 182-188.

[8] LI Xiaona, SHEN Xing-lai, Xue Xue, et al. Power quality disturbance identification based on improved HHT and decision tree [J]. Journal of Electric Power Construction, 2017, 38(02): 114-121.

[9] Wang Jiye, Zhu Xinyan, Zhao Guang, et al. Research on optimal outage model based on deep artificial neural network and GIS data [J]. Power System Protection and Control, 2019, 47(16): 58-63.

[10] Guo Yunfeng, Yang Xiaomei. Power quality disturbance signal classification based on SVM [J]. Computer Applications and Software, 2022, 39(07): 95-100+120.

[11] Li Zuming, Lv Qianyun, Chen Nuo, et al. Power quality complex disturbance recognition based on chaotic integrated decision Tree [J]. Power System Protection and Control, 2021, 49(21): 18-27.

[12] Zhang Minglong, Zhang Zhenyu, Luo Xiang, et al. Research on Hybrid disturbance waveform identification algorithm based on Multi-core support Vector Machine [J]. Power System Protection and Control, 2022, 50(15): 43-49.

[13] Wang Renming, Wang Hongyang, Zhang Yunning, et al. Hybrid power quality disturbance identification Method based on segmental improved S-transform and Random forest [J]. Protection and Control of Power Systems, 2020, 48(07): 19-28.

[14] XU Yanchun, Fan Shirong, Tan Chao, et al. Power Quality Disturbance detection and classification of high permeability active distribution Network based on improved EWT-CMPE [J]. Power Grid Technology, 2020, 44(10): 3991-4005.

[15] YAO Li, Sun Jianjun, Ma Chenbo. Fault diagnosis Method of Rolling Bearing Based on Gram Angle Field and CNN-RNN [J]. Journal of Bearings, 2022(02): 61-67.

[16] Yang Wei, Pu Caixia, Yang Kun, et al. Short-term Fault prediction Method of Transformer Based on CNN-GRU Combined neural Network [J]. Power Systems Protection and Control, 2022, 50(06): 107-116.

[17] Xu Dong, Yang Guan, Liu Xiaoming, et al. Small sample image classification based on adaptive feature fusion and transformation [J]. Computer Engineering and Applications, 2022, 58(24): 223-232.

[18] Song Tiewei, Shi Weifeng, Bi Zong, et al. Power quality Disturbance Identification of Marine Power System based on 2D-ResNet [J]. Power System Protection and Control, 2022, 50(10): 94-103.